

University of Louisville

ThinkIR: The University of Louisville's Institutional Repository

Faculty Scholarship

7-2005

Verhaltensbeurteilung durch Gleichaltrige. Uebersetzung und Evaluierung der Adjustment Scales for Sociometric Evaluation of Secondary-School Students (ASSESS-D) [Peer rating of behavior. Translation and evaluation of the Adjustment Scales for Sociometric Evaluation of Secondary-School Students (ASSESS-D)].

Patrick Pössel
University of Louisville

Ulrike von Delleman
Universität Tübingen

Martin Hautzinger
Universität Tübingen

Follow this and additional works at: <https://ir.library.louisville.edu/faculty>



Part of the [Counseling Psychology Commons](#)

Original Publication Information

Pössel, Patrick, Ulrike von Delleman and Martin Hautzinger. "Verhaltensbeurteilung durch Gleichaltrige. Uebersetzung und Evaluierung der Adjustment Scales for Sociometric Evaluation of Secondary-School Students (ASSESS-D). [Peer rating of behavior. Translation and evaluation of the Adjustment Scales for Sociometric Evaluation of Secondary-School Students (ASSESS-D).]" 2005. *Zeitschrift fuer Entwicklungspsychologie und Paedagogische Psychologie* 37: 135 - 143.

This Article is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. For more information, please contact thinkir@louisville.edu.

Kolumnentitel: VERHALTENSBEURTEILUNG DURCH GLEICHALTRIGE

Verhaltensbeurteilung durch Gleichaltrige. Übersetzung und
Evaluierung der Adjustment Scales for Sociometric
Evaluation of Secondary-School Students (ASSESS-D)

Patrick Pössel, Ulrike von Delleman und Martin Hautzinger
Universität Tübingen

Zusammenfassung

Beurteilungen durch Gleichaltrige gelten als valide und reliable Informationsquelle, dennoch gibt es im deutschen Sprachraum kein methodisch bewährtes Erhebungsinstrument zur Beurteilung des Verhaltens von Jugendlichen durch Gleichaltrige. Um diese Lücke zu füllen wurden die „Adjustment Scales for Sociometric Evaluation of Secondary-School Students“ (ASSESS) von Prinz, Swan, Liebert, Weintraub und Neale (1978) übersetzt und evaluiert. Die an 301 Jugendlichen faktorenanalytisch ermittelten fünf Skalen klären 66.77 % der Gesamtvarianz auf. Die Internen Konsistenzen (Cronbachs Alpha) der 34 Items umfassende ASSESS-D liegen zwischen $\alpha = .81$ und $.95$ und die Retest-Reliabilität über drei Wochen liegt zwischen $r = .84$ und $.95$. Die Korrelationen mit externalisierenden Störungen sind überwiegend erwartungskonform, während sich keine bedeutsamen Zusammenhänge mit Internalisierungsstörungen zeigen. Varianzanalysen zeigten die erwarteten Unterschiede zwischen Geschlecht, Alter und Schultyp. Mit der deutschen Version der ASSESS liegt nach den bisherigen Daten ein ökonomisches und reliables Instrument zur Erfassung von Verhalten Jugendlicher aus der Sicht Gleichaltriger vor.

Schlagwörter: Beurteilung durch Gleichaltrige, psychometrische Eigenschaften, Internalisierungs- und Externalisierungsstörungen.

Peerrating of behavior. Translation and evaluation of the
Adjustment Scales for Sociometric Evaluation of Secondary-
School Students (ASSESS-D)

Abstract

Peers are valide and reliable informants. Hence there is no evaluated german questionnaire to assess the ratings of adolescent behavior by peers. Therefore we translated the "Adjustment Scales for Sociometric Evaluation of Secondary-School Students" (ASSESS) of Prinz, Swan, Liebert, Weintraub, and Neale (1978) and evaluated them on a sample of 301 youth. Five factors (accounted variance 66.77 %) were extracted. The internal consistencies (Cronbachs Alpha) of the 34 item enclosing ASSESS-G ranges between $\alpha = .81$ and $.95$ and three weeks retest reliability ranges between $r = .75$ and $.95$. Correlations with externalising problems are predominantly in agreement with the expectations, while no important connections to internalising problems could be observed. Analyses of variance showed the expected differences between gender, age and type of school. The German version of the ASSESS seems to be an economical and reliable instrument for recording behavior of adolescents by peers.

Keywords: Peer rating, psychometric properties, internalising and externalising Problems.

Theorie

In der Literatur wird die Bedeutung von Gleichaltrigen als Informanten zur Identifikation von Kindern und Jugendlichen mit psychosozialen Schwierigkeiten und Verhaltensauffälligkeiten diskutiert (Scheithauer, Hayer & Petermann, 2003). So gelten sie als unverzichtbare Informanten, da sie gleiche bzw. ähnliche Lebensräume wie die zu beschreibenden Jugendlichen haben und ähnliche Erfahrungen teilen. Aufgrund dessen verfügen Gleichaltrige über ein ausgeprägtes Wissen bezüglich der sozialen Normen und Regeln der jeweiligen Subkultur (Newcomb, Bukowski & Pattee, 1993; Realmuto, August, Seiler & Pessoa-Brandao, 1997), so dass sie beobachtbares Verhalten besser einschätzen können als erwachsene Fremdbeurteiler. Darüber hinaus können sie auf Grund der häufigen Kontakte selten auftretendes Verhalten besser beobachten als Eltern oder Lehrer (Perry, Kusel & Perry, 1988). Gestützt wird die Bedeutung von Urteilen durch Gleichaltrige auch durch die Metaanalyse von Achenbach, McConaughy und Howell (1987). In dieser Metaanalyse zeigt sich, dass die Beurteilerübereinstimmung beim Urteil Gleichaltriger mit $r = .73$ ähnlich hoch ist wie die Korrelation zwischen verschiedenen Selbsturteilen ($r = .74$), jedoch deutlich über den Beurteilerübereinstimmungen von Eltern ($r = .59$), Lehrern ($r = .64$), psychologischen Experten ($r = .54$) und geschulten Beobachtern ($r = .57$) liegt. Weiterhin hat sich

die Beurteilung von Gleichaltrigen durch Gruppen als zuverlässig erwiesen, da die Informationen für jeden Schüler über die jeweils beurteilenden Gleichaltrigen gemittelt werden (Scheithauer et al., 2003). Auf diese Weise wird beispielsweise der Freundschaftsbias minimiert, der normalerweise auftritt, wenn nur der beste Freund oder die beste Freundin des zu beurteilenden Jugendlichen befragt wird (Kane & Lawler, 1978). Weiterhin sind Beurteilungen durch Gleichaltrige relativ frei vom Reaktionsbias (Kane & Lawler, 1978), da dieser Bias eine Reihe von widersprüchlichen Tendenzen umfasst (z. B. Ja-Sage-Tendenz vs. soziale Erwünschtheit), die nicht bei allen Beurteilern gleichermaßen ausgeprägt sind. Auf diese Weise kommt der Reaktionsbias nicht so stark zum Tragen wie bei nur einem Beurteiler (z. B. Selbsturteil, Lehrer, Freund), weshalb sich die statistische Reliabilität erhöht (Perry et al., 1988). Weiterhin sind Beurteilungen durch Gleichaltrige nicht nur sensitiver als Selbstbeurteilungen (Lefkowitz & Tesiny, 1980), sondern auch stabiler als Selbst- und Lehrerurteile (Tesiny & Lefkowitz, 1982).

Jedoch hat die Erfassung von Verhalten durch Gleichaltrige nicht nur Vorteile. So werden auch Gleichaltrige in ihrem Urteil durch ihre Einstellungen dem zu beurteilenden Jugendlichen gegenüber und durch die Subkultur, welcher der zu Beurteilende angehört, beeinflusst (Scheithauer et al., 2003). Dieser so genannte

Reputationsbias ist besonders wichtig, wenn sich die Beurteilenden und der zu Beurteilende lange kennen, wenn es sich um enge Freunde handelt, sowie wenn die Subkultur des zu beurteilenden Jugendlichen klar definiert und leicht zu erkennen ist (z. B. Skater, Ausländer; Inderbitzen, 1994).

Weiterhin unterliegen Peerbeurteilungsverfahren durch größere Gruppen dem Problem der Intergruppen-Reliabilität, d. h. unterschiedliche Gruppen von Gleichaltrigen beurteilen das gleiche Verhalten zum Teil unterschiedlich. Hierbei ist jedoch anzumerken, dass Urteile von Gleichaltrigen - ebenso wie alle anderen Informationen aus einer Quelle - eine wichtige, aber nicht die einzige Informationsquelle darstellen sollten.

Gegen Urteile von Gleichaltrigen durch größere Gruppen kann weiterhin eingewendet werden, dass möglicherweise nur bei solchen Jugendlichen valide Urteile möglich sind, die relativ eindeutig bestimmtes (z. B. aggressives, schüchternes) Verhalten zeigen bzw. denen von allen Gleichaltrigen eindeutig bestimmtes Verhalten zugeordnet wird. Fraglich ist hingegen, wie valide die gemittelten Urteile bei allen anderen Jugendlichen sind. Um dies zu klären, ist es notwendig, die Varianz dieser beiden Gruppen (Jugendliche mit eindeutigen vs. wechselnden Beurteilungen durch die Gleichaltrigen) miteinander zu vergleichen. Ist die Varianzen in der Gruppe mit wechselnden Beurteilungen durch die Gleichaltrigen signifikant größer als in der

Gruppe mit eindeutigen Beurteilungen, würde dies auf eine nur eingeschränkte Einsetzbarkeit von Peerbeurteilungsmaßen hinweisen.

Trotz dieser Probleme wird die Bedeutung der Beurteilung Jugendlicher durch Gleichaltrige allgemein akzeptiert. Dennoch gibt es im deutschsprachigen Raum kein methodisch bewährtes Erhebungsinstrument zur Beurteilung des Verhaltens von Jugendlichen durch Gleichaltrige. Ein Beispiel für Peerbeurteilungsverfahren sind die „Adjustment Scales for Sociometric Evaluation of Secondary-School Students“ (ASSESS) von Prinz, Swan, Liebert, Weintraub und Neale (1978). Hierbei handelt es sich um einen Fragebogen zur Erfassung von Beurteilungen der Verhaltensbereiche Aggression, Rückzug, Angst, Soziale Kompetenz und Schulische Schwierigkeiten durch Gleichaltrige, bei dem Jugendliche mit insgesamt 41 Beschreibungen in einer Person x Item- Matrix angeben sollen, auf welche(n) ihrer Altersgenossen die Beschreibungen zutreffen. In einer Evaluationsstudie an 277 Mädchen und Jungen der Klassenstufen 9 bis 12 reichten die Trennschärfen der amerikanischen ASSESS von .60 bis .93 (Prinz et al., 1978). Die Mehrzahl der Skalen-Interkorrelation bewegt sich zwischen -.30 und .30, so dass die Skalen keinen großen Anteil gemeinsamer Varianz aufweisen. Die interne Konsistenz liegt für die Subskalen zwischen $\alpha = .84$ und

.95 und die dreiwöchige Retest-Reliabilität zwischen $r = .72$ und $.91$ (Prinz et al., 1978).

Die Übersetzung und Evaluation der Adjustment Scales for Sociometric Evaluation of Secondary-School Students (ASSESS) von Prinz et al. (1978) diente dem Ziel ein reliables, valides und zeitökonomisches Maß zur Beurteilung von Verhalten von Jugendlichen durch Gleichaltrige für den deutschsprachigen Raum zur Verfügung zu stellen.

Da unklar ist, ob Urteile von Gleichaltrigen durch größere Gruppen auch bei Jugendlichen valide sind, die nicht eindeutig ein bestimmtes (z. B. aggressives, schüchternes) Verhalten zeigen, ist es ein weiteres Ziel, die Validität der deutschen Version der ASSESS bei Schülern mit wechselnden Beurteilungen durch Gleichaltrige zu überprüfen.

Weiterhin soll der Zusammenhang zwischen Geschlecht, Alter und Schultyp einerseits und den mit den ASSESS erhobenen Verhaltensweisen andererseits untersucht werden. Nach bisherigen Studien steigt Rückzugsverhalten mit zunehmendem Alter sowohl im Selbst- als auch im Elternurteil an. Während sich im Elternurteil jedoch kein Geschlechtseffekt findet, berichten Mädchen im Selbsturteil mehr Rückzugsverhalten als Jungen (Döpfner et al., 1997). Ähnlich verhält es sich nach Döpfner et al. (1997) mit Angst. Nur zeigt sich hier hauptsächlich im Selbsturteil der Mädchen ein Anstieg, während Jungen und Eltern keine

Veränderungen berichten. Bei aggressivem Verhalten findet sich im Selbst- und Elternurteil einen Anstieg bis zum Alter von etwa 14 Jahren, der von einem Absinken von Aggressionen bis zum Alter von 18 Jahren gefolgt wird (Döpfner et al., 1997; Lindeman, Harakka & Keltikangas-Järvinen, 1997). Ein Geschlechtseffekt ist nur im Eltern-, nicht aber im Selbsturteil nachweisbar. Eltern schätzen Mädchen als weniger aggressiv ein als Jungen (Döpfner et al., 1997). Weiterhin zeigen sich ein Anstieg der sozialen Kompetenz und eine Abnahme schulischer Probleme mit zunehmendem Alter. Unabhängig hiervon werden Mädchen mehr soziale Kompetenz und weniger schulische Probleme zugeschrieben als Jungen (Prinz et al., 1978).

Der Zusammenhang zwischen Schultyp und den mit den ASSESS erhobenen Verhaltensweisen wurde unseres Wissens nach bisher kaum berücksichtigt, weshalb diese Zusammenhänge in der vorliegenden Studie untersucht werden sollten. Insgesamt zeigt sich aber, dass in Hauptschulen mehr abweichendes Verhalten auftritt, als in den anderen beiden Schultypen, welche sich wiederum kaum voneinander unterscheiden (Remschmidt & Walter, 1990). Allerdings sind uns für die einzelnen durch die ASSESS erfassten Verhaltensweisen keine Ergebnisse bekannt, so dass diese Analysen explorativ sind.

Schließlich wird erwartet, dass die ASSESS-D Subskalen „Aggression“ und „Schulische Probleme“ mit

externalisierenden Störungen, die Subskalen der ASSESS-D „Rückzug“, „Angst“ und „Schulische Probleme“ mit Internalisierungsstörungen, sowie die ASSESS-D Subskalen „Aggression“ und „Schulische Probleme“ mit anderen Problemen wie Aufmerksamkeitsproblemen positiv korrelieren. Weiterhin werden zwischen den ASSESS-D Subskalen „Rückzug“, „Angst“ und „Soziale Kompetenz“ und zwischen externalisierenden Störungen, sowie zwischen der der ASSESS-D Subskala „Soziale Kompetenz“ und Internalisierungsstörungen signifikante negative Korrelationen erwartet.

Methodik

Stichprobe

Die Übersetzung und Validierung der ASSESS erfolgte an einer Stichprobe von Jugendlichen aus einer Schülerpopulation der Region Tübingen ($\underline{n} = 314$). Das durchschnittliche Alter der Jugendlichen lag bei 14.27 ($\underline{SD} = 1.01$) mit dem Median von 14 Jahren und einem Range von 13 bis 17 Jahren. Der Anteil der Mädchen lag bei 51 % ($\underline{n} = 160$), wobei Alter und Geschlecht nicht signifikant miteinander korrelierten ($\underline{r} = -.01$; $\underline{p} = .861$). Die Datenerhebungen erfolgten im Rahmen des normalen Schulunterrichtes, wobei die Jugendlichen an zwei Datenerhebungen im Abstand von 15-22 Tagen ($\underline{M} = 18.5$; $\underline{SD} = 2.74$) teilnahmen. Die Schüler und Schülerinnen verteilten

sich auf die Klassenstufen 8 ($\underline{n} = 131$), 9 ($\underline{n} = 122$) und 10 ($\underline{n} = 61$) und entstammten allen drei Schultypen der Regelschule (Gymnasiasten: $\underline{n} = 117$; Realschüler: $\underline{n} = 104$; Hauptschüler: $\underline{n} = 93$).

Um dem Datenschutz bei namentlicher Verhaltenseinschätzung sicher zustellen, wurden die Jugendlichen zu Beginn der Erhebung gebeten ihre Namen in zufälliger Reihenfolge an die Tafel zu schreiben. Die Namen wurden dann in der so entstandenen Reihenfolge durchnummeriert, so dass jedem Schüler eine Zahl zugeordnet wurde, die allen im Raum Anwesenden bekannt und ständig präsent war. Diese Zuordnung von Name und Zahl diente der Kodierung sowohl des eigenen Fragebogens, als auch der Beurteilung der Klassenkameraden. Auf diese Weise fand noch im Klassenzimmer eine Trennung von Namen und Daten statt. Diese Liste wurde von einem Projektmitarbeiter notiert und zur zweiten Messung wieder an die Tafel geschrieben. Unmittelbar nach Abschluss der zweiten Messung in der jeweiligen Klasse wurde diese Liste vernichtet.

Bei der Auswertung des ersten Messzeitpunktes konnten die Fragebögen von insgesamt 11 Schülern (3.5 %) nicht verwendet werden, da diese nicht alle Items korrekt ausfüllten. Bei der Auswertung der Daten des zweiten Messzeitpunktes konnten die Fragebögen von 37 Schülern (11.8 %) nicht in die Auswertung mit einbezogen werden, da 31 nicht erschienen und 6 weitere Schüler die Fragebögen

nicht korrekt ausfüllten. Demnach gehen zum zweiten Messzeitpunkt nur die Daten von 277 (88.2 %) Jugendlichen in die Datenanalysen ein.

Instrumente

Bei den Adjustment Scales for Sociometric Evaluation of Secondary-School Students (ASSESS) handelt es sich um einen Beurteilungsbogen für Schülerinnen und Schülern zur Erfassung von Verhalten von Gleichaltrigen. Die deutschsprachige Vorform der ASSESS besteht aus 41 aus dem amerikanischen übersetzten Items, die in einer Item x Person-Matrix dargeboten werden. Diese Vorform der ASSESS wurde erzeugt, in dem das amerikanischsprachige Original vom Erstautor aus dem Englischen ins Deutsche übersetzt und diese deutsche Version anschließend von einer amerikanischsprachigen Muttersprachlerin wiederum ins Englische zurückübersetzt wurde. Wie im amerikanischen Original werden die Schüler angewiesen, bei jeder einzelnen Beschreibung schriftlich anzugeben, auf welche(n) Mitschüler die jeweilige Beschreibung zutrifft, wobei „Trifft nicht zu“ mit Null und „Trifft zu“ mit Eins gewertet wurde.

Zur Bestimmung der Kriteriumsvalidität wurde neben den ASSESS der „Youth Self Report“ (YSR; Arbeitsgruppe Deutsche Child Behavior Checklist, 1998) vorgelegt. Beim YSR handelt es sich um einen Selbstbeurteilungsfragebogen zur Erfassung

von Verhalten und psychopathologischen Symptomen bei Jugendlichen im Alter von 11 bis 18 Jahren. Der in der vorliegenden Studie verwendete Teil des YSR besteht aus 119 Items, von denen 103 Items Verhaltensauffälligkeiten, emotionale Auffälligkeiten, körperliche Beschwerden, sowie sozial unerwünschte Verhaltensweisen erfassen und 16 Items sozial erwünschte Verhaltensweisen beschreiben. Diese Skalen lassen sich den Faktoren "Internalisierende Störungen", „Externalisierende Störungen“ und „Andere Probleme“ zuordnen. Cronbachs Alpha für die Gesamtauffälligkeit des YSR liegt für Jungen bei $\alpha = .92$ und für Mädchen bei $\alpha = .91$ (Arbeitsgruppe Deutsche Child Behavior Checklist, 1998).

Datenanalyse

Für jeden Schüler wurden die Angaben der Klassenkameraden gemittelt. Diese Mittelwerte dienten als Grundlage für alle Analysen. So auch zur Überprüfung, ob die ASSESS-D bei allen Jugendlichen valide sind, unabhängig davon, ob die zu beurteilenden Jugendlichen eindeutig ein Verhaltensmuster oder wechselndes Verhalten zeigen. Hierzu wurde die Varianz beider Gruppen mit Brown-Forsythe-2V-Tests miteinander verglichen. Der Gruppe von Schülern mit eindeutigem Verhalten wurden alle Jugendlichen zugeordnet, die auf jeweils einer Skala der ASSESS-D mindestens eine Standardabweichung über dem Mittelwert der Gesamtgruppe und

in allen anderen Skalen unter dem Mittelwert der Gesamtgruppe liegen ($\underline{n} = 33$). Auf diese Weise wurde eine Extremgruppe mit eindeutigen Zuordnungen ausschließlich eines Verhaltensmusters zu dem jeweiligen Schüler durch die Gleichaltrigen erreicht. Alle anderen Schüler wurden der Gruppe zu geordnet, deren Verhalten wechselnd ist ($\underline{n} = 268$).

Um den Zusammenhang zwischen Geschlecht, Alter und Schultyp einerseits und den mit den ASSESS erhobenen Verhaltensweisen andererseits zu untersuchen, wurde eine drei-faktorielle Varianzanalyse (MANOVA) mit den Faktoren Alter (≤ 14 Jahre vs. > 14 Jahre), Geschlecht (Jungen vs. Mädchen) und Schultyp (Hauptschule vs. Realschule vs. Gymnasium) durchgeführt. Als Anschluss-tests wurden Bonferroni-adjustierte t-Tests berechnet.

Alle Analysen wurden mit dem Statistikprogramm SPSS 11.0 für Windows durchgeführt.

Ergebnisse

Itemanalyse

Alle 41 Items der deutschen Übersetzung des ASSESS weisen eine Itemschwierigkeit (Mittelwert der Angaben aller Schüler über alle anderen Schüler in diesem Item) zwischen .05 und .95 auf und genügen damit den von Prinz et al. (1978) aufgestellten psychometrischen Anforderung (Tabelle 1).

Weiterhin wurde eine Hauptkomponentenanalyse durchgeführt, nach der den ASSESS fünf Faktoren mit einem Eigenwert von mindestens eins (Kaiser-Kriterium) zu Grunde liegen. Die Items wurden demjenigen Faktor zugeordnet, auf dem sie mit mindestens .40 laden. Die Items 11, 19, 20, 21, 24 und 29 weisen keine Differenz von $\leq .10$ zwischen höchster und zweithöchster Ladung auf und wurden daher aus dem Itempool entfernt (Bortz & Döring, 2002). Die fünf Faktoren klären nach der Varimax-Rotation insgesamt 66.77 % der Gesamtvarianzvarianz auf. Die Zuordnung der Items zu den Faktoren, die Faktorladungen der Items, die Eigenwerte und die aufgeklärte Varianz sind in Tabelle 2 dargestellt.

Vier der fünf Faktoren entsprechen den Skalen in der amerikanischen Originalversion (Prinz et al., 1978). Lediglich für die Skala Angst der amerikanischen Originalfassung findet sich keine Entsprechung in den ASSESS-D, da einige Items aufgrund mangelhafter psychometrischer Eigenschaften gestrichen wurden. Aufgrund des hieraus folgenden anderen Inhaltes wurde diese Skala in der deutschen Fassung in „Empfindsamkeit“ umbenannt.

Die korrigierte Trennschärfe der verbleibenden Items mit den Skalen liegen zwischen $\underline{rit} = .54$ und $\underline{rit} = .88$. Lediglich Item 18 liegt mit einer korrigierten Trennschärfe von $\underline{rit} = .44$ unter der von Diehl und Kohr (1989) geforderten korrigierten Trennschärfe von \underline{rit} Items $\geq .50$ und wurde dementsprechend ausgeschlossen (Tabelle 1).

Reliabilität

Die Interne Konsistenz (Cronbachs Alpha) ist sowohl für den Gesamttest ($\alpha = .81$), als auch für die einzelnen Subskalen mit Werten zwischen $\alpha = .84$ und $\alpha = .95$ zufrieden stellend bis gut (Bortz & Döring, 2002). Die Retest-Reliabilität über drei Wochen reicht für die ASSESS-D von $r = .84$ bis zu $r = .95$ und stellt sich damit ebenfalls zufrieden stellend bis gut dar. Ausschließlich die Interkorrelationen zwischen „Aggression“ und „Schulische Probleme“ ($r = .63$) wie auch zwischen „Rückzug“ und „Empfindsamkeit“ ($r = .57$) sind substantiell (Lutz, 1989), was in Übereinstimmung mit den Erwartungen steht (Tabelle 3). Dieses Muster ist identisch, wenn die Interkorrelationen für beide Geschlechter getrennt berechnet werden.

Validität

Zur Ermittlung der Kriteriumsvalidität wurden die Daten der ASSESS-D Skalen mit den Youth Self Report (YSR; Arbeitsgruppe Deutsche Child Behavior Checklist, 1998) Breitbandskalen „Externalisierende Störungen“, „Internalisierende Störungen“ und „Andere Probleme“ korreliert (Tabelle 4).

Erwartungsgemäß ergaben sich signifikante positive Korrelationen zwischen der YSR Breitbandskala

„Externalisierende Störungen“ und den ASSESS-D Subskalen „Aggression“ und „Schulische Probleme“, sowie eine signifikant negative Korrelation zur ASSESS-D Subskala „Rückzug“. Kein Zusammenhang zeigt sich zwischen „Externalisierende Störungen“ und der ASSESS-D Skala „Soziale Kompetenz“. Auch korreliert keine der ASSESS-D Skalen mit der YSR Skala „Internalisierende Störungen“. Jedoch zeigen sich zwischen der YSR Skala „Andere Probleme“ und den Skalen „Aggression“ und „Schulische Probleme“ erwartungskonform positive Korrelationen. Die deutsche Skala „Empfindsamkeit“ zeigt keine signifikanten Zusammenhänge mit den YSR Breitbandskalen.

Weiterhin wurde die Varianz der Beurteilung von Schülern mit eindeutigem Verhalten mit der von Jugendlichen mit wechselhaftem Verhalten auf den Skalen der ASSESS-D miteinander verglichen, um zu testen, ob die ASSESS-D bei allen Jugendlichen valide Ergebnisse ergeben. Hierbei zeigten sich auf keiner Skala signifikante Unterschiede beider Varianzen (Aggression: $t(299) = .91$; $p > .20$; Rückzug: $t(299) = .27$; $p > .20$; Empfindsamkeit: $t(299) = .17$; $p > .20$; Soziale Kompetenz: $t(299) = .00$; $p > .20$; Schulische Probleme: $t(299) = .48$; $p > .20$), so dass die ASSESS-D auch bei Jugendlichen einsetzbar sind, die wechselndes Verhalten zeigen.

Einfluss von Alter, Geschlecht und Schultyp auf die Ausprägung auf den ASSESS-D Skalen

Um geschlechtsspezifische, entwicklungsbedingte und schultypabhängige Unterschiede in den mit den ASSESS-D gemessenen Verhaltensweisen zu ermitteln, wurde eine dreifaktorielle Varianzanalyse (MANOVA) mit den Faktoren Alter (≤ 14 Jahre vs. > 14 Jahre), Geschlecht (Jungen vs. Mädchen) und Schultyp (Hauptschule vs. Realschule vs. Gymnasium) durchgeführt. Hierbei zeigten sich Haupteffekte für das Geschlecht in den Skalen „Aggression“ ($F(1/287) = 23.68$, $p < .001$), „Empfindsamkeit“ ($F(1/287) = 17.28$, $p < .001$), „Soziale Kompetenz“ ($F(1/287) = 17.40$, $p < .001$) und „Schulische Schwierigkeiten“ ($F(1/287) = 22.08$, $p < .001$) und für den Schultyp in der Skala „Aggression“ ($F(2/287) = 4.53$, $p = .012$). Weiterhin wurde die Interaktion zwischen Alter und Schultyp bei der Skala „Soziale Kompetenz“ signifikant ($F(2/287) = 4.02$, $p = .019$). Jungen erhalten hierbei höhere Werte in den Skalen „Aggression“ und „Schulische Schwierigkeiten“ als Mädchen, während es in den Skalen „Empfindsamkeit“ und „Soziale Kompetenz“ umgedreht ist. Weiterhin werden Hauptschüler von ihren Altersgenossen als aggressiver eingeschätzt als dies bei Gymnasiasten der Fall ist ($t(197) = 2.08$; $p = .002$). Die Interaktion von Alter und Schultyp schließlich ist darauf zurückzuführen, dass die jüngeren Gymnasiasten als sozial kompetenter eingeschätzt werden als ihre Altersgenossen in

Hauptschulen ($t(131) = 8.90; p = .005$). Dieser Effekt zeigt sich nicht mehr bei den älteren Schülern.

Diskussion

Die deutsche Form der Adjustment Scales for Sociometric Evaluation of Secondary-School Students (ASSESS-D) besteht aus den Skalen „Aggression“ (10 Items), „Rückzug“ (6 Items) und „Empfindsamkeit“ (2 Items), „Soziale Kompetenz“ (9 Items) und „Schulische Probleme“ (7 Items). Insgesamt klären die fünf Faktoren 66.77% der Varianz auf. Vier der fünf Skalen zeigen eine solch hohe inhaltliche Übereinstimmung mit den Skalen von Prinz et al. (1978), dass deren Benennung übernommen wurde. Lediglich die Skala „Angst“ der amerikanischen Originalfassung wurde in die nur zwei Items umfassende Skala „Empfindsamkeit“ umbenannt, da einige Items des Originals aufgrund mangelhafter psychometrischer Eigenschaften nicht berücksichtigt wurden und sich hieraus eine inhaltliche Veränderung ergab. Sowohl die Interne Konsistenz für die ASSESS-D (Cronbachs Alpha: $\alpha = .81$) und deren Skalen (Cronbachs Alpha: $\alpha = .84 - .95$) als auch die Retest-Reliabilitäten ($r = .84 - .95$) sind zufrieden stellend hoch. Insgesamt weisen die ASSESS-D damit vergleichbare oder sogar bessere psychometrische Eigenschaften auf als das amerikanische Original. Dies gilt auch für die nur zwei Items umfassende Skala „Empfindsamkeit“.

Weiterhin sind ausschließlich die Korrelationen zwischen „Aggression“ und „Schulische Probleme“, sowie zwischen „Rückzug“ und „Empfindsamkeit“ substantiell (Lutz, 1989). Dieses Muster zeigt sich sowohl, wenn die Interkorrelationen mit beiden Geschlechtern gemeinsam als auch wenn diese getrennt berechnet werden. Dies stützt die Relevanz aller fünf extrahierten Skalen, was auch die sehr kurze Skala „Empfindsamkeit“ mit einschließt. Dennoch kann die Korrelation zwischen „Aggression“ und „Schulische Probleme“ so interpretiert werden, dass sie eigenständige Anteile von oppositionellem Verhalten abbilden. Der substantielle Zusammenhang zwischen „Rückzug“ und „Empfindsamkeit“ kann ebenso als „Sensibilität“ verstanden werden. Eine starke Ausprägung beider Verhaltensmuster stellt im schulischen Kontext nicht nur für die Mitschüler, sondern insbesondere für die Lehrer eine Quelle erheblicher Probleme dar.

Auch die Korrelationen zwischen Externalisierungsstörungen einerseits und Aggression, Rückzug und schulischen Problemen andererseits, sowie zwischen anderen Problemen (z. B. Aufmerksamkeitsprobleme) und Aggression und schulischen Problemen entsprechen den Erwartungen. Der fehlende Zusammenhang zwischen von Gleichaltrigen beurteiltem Verhalten (z. B. Rückzug, soziale Kompetenz und schulische Probleme) und selbstberichteten Internalisierungsstörungen entspricht

zwar nicht den Erwartungen, ist aber in Übereinstimmung mit früheren Studien, die keine oder geringe Zusammenhänge zwischen selbst- und fremdbeurteilen

Internalisierungsstörungen aufzeigen. Die wird darauf zurückgeführt, dass Internalisierungsstörungen Symptome umfassen, die nicht leicht von außen zu beobachten sind und sich nicht in beobachtbarem Verhalten niederschlagen müssen (z. B. Schuldgefühle, Hilflosigkeit oder Gefühle der Wertlosigkeit, Albayrak-Kaymak, 1999; Reynolds, 1998).

Weiterhin zeigte der Vergleich der Varianzen von Schülern mit eindeutigen und wechselnden Beurteilungen keine bedeutsamen Differenzen. Dieses Ergebnis entkräftet den Einwand, dass Urteile von Gleichaltrigen durch größere Gruppen möglicherweise nur bei solchen Jugendlichen valide einsetzbar sind, denen relativ eindeutig bestimmtes (z. B. aggressives, schüchternes) Verhalten durch ihre Altersgenossen zugeordnet wird, nicht jedoch bei Jugendlichen mit wechselnden Beurteilungen.

Bei den Analysen zum Zusammenhang zwischen Alter, Geschlecht und Schultyp einerseits und den mit den ASSESS-D erfassten Verhaltensweisen andererseits entsprechen die Geschlechtseffekte den Erwartungen und der Literatur (Döpfner et al., 1997; Prinz et al., 1978). So wird Mädchen von ihren Altersgenossen eher internalisierendes Verhalten und soziale Kompetenz zugeschrieben, während Jungen eher externalisierende Verhaltensweisen zugeordnet werden. Dass

Alter hingegen kaum eine Rolle spielt, kann wahrscheinlich auf den geringen Range von 13 bis 17 Jahren im Vergleich zu bisherigen Studien (Döpfner et al., 1997: 4 - 18 Jahre; Lindeman, Harakka & Keltikangas-Järvinen, 1997: 11 -18 Jahre) zurückgeführt werden. Interessant ist weiterhin, dass der Schultyp ausschließlich bei den externalisierenden Verhaltensweisen Aggression und soziale Kompetenz bedeutsam zu sein scheint, wobei der Schultyp bei sozialer Kompetenz darüber hinaus mit zunehmendem Alter an Bedeutung verliert.

Einschränkend ist, dass bisher keine Normierungsdaten vorliegen. Unabhängig hiervon kann die ASSESS-D dennoch zur Evaluation von Trainingsprogrammen beispielsweise von sozialer Kompetenz (Frankel, Myatt & Cantwell, 1995) eingesetzt werden. Auch entwickeln Kinder von Eltern mit psychischen Störungen gehäuft Verhaltensauffälligkeiten. Diese können sowohl in ihrem Verlauf als auch im Vergleich zu Kindern mit unauffälligen Eltern mit einem Peerbeurteilungsverfahren erfasst werden. Urteile durch Gleichaltrige scheinen in diesen Bereichen besonders geeignete Informationsquellen zu sein, da das Zielverhalten häufig in Interaktionen mit Gleichaltrigen auftritt.

Beim Einsatz der ASSESS-D ist allerdings zu berücksichtigen, dass Jugendliche über Gleichaltrige namentliche Auskunft geben sollen. Um dies zu ermöglichen, erscheint nicht nur eine genaue Aufklärung der Jugendlichen und ihrer Eltern, sondern auch die Garantie der Anonymität

und Vertraulichkeit der Daten unbedingt notwendig. Auch die relativ hohe Drop-out Rate von fast 12 % kann mit diesem Problem von Peerbeurteilungsmaßen in Verbindung stehen. Allerdings gibt es weder in unserer Studie, noch in der vorliegenden Literatur objektivierbare Hinweise auf Probleme bei Beurteilungen durch Gleichaltrige (z. B. Merrell, 2001). Deshalb gehen wir von einer breiten Einsetzbarkeit der ASSESS-D aus, solange diesen von uns genannten Punkten Rechnung getragen wird.

Insgesamt liegt nach den bisherigen Daten mit den ASSESS-D ein Instrument vor, das die reliable und zeitökonomische Erfassung des Verhaltens von Jugendlichen aus Sicht ihrer Altersgenossen erlaubt. Von besonderer Bedeutung sind hierbei die bedeutsamen Zusammenhänge zwischen durch Gleichaltrige beurteilter Aggression, Rückzug und schulischen Problemen einerseits und selbstberichteten Externalisierungsstörungen andererseits. Weiterhin scheinen die ASSESS-D das Verhalten von allen Schülern valide zu erfassen und können geschlechts-, alters- und schultypabhängige Differenzen erfassen.

Literatur

Achenbach, T.M., McConaughy, S.H. & Howell, C.T. (1987). Child/Adolescent Behavioral and Emotional Problems: Implications of Cross-Informant Correlations for Situational Specificity. Psychological Bulletin, 101, 213-232.

Albayrak-Kaymak, D. (1999). Internalizing or Externalizing: Screening for both Problem Youth. International Journal for the Advancement of Counselling, 21, 125 - 137.

Arbeitsgruppe Deutsche Child Behavior Checklist: Fragebogen für Jugendliche; deutsche Bearbeitung des Youth Self-Report (YSR) der Child Behavior Checklist. Arbeitsgruppe Kinder-, Jugend- und Familiendiagnostik (KJFD), Köln 1998.

Diehl, J. M. & Kohr, H. U. (1989). Deskriptive Statistik. Eschborn: Verlag Dietmar Klotz.

Döpfner, M., Plück, J., Berner, W., Fegert, J. M., Huss, M., Lenz, K., Schmeck, K., Lehmkuhl, U., Poustka, F. & Lehmkuhl, G. (1997). Psychische Auffälligkeiten von Kindern und Jugendlichen in Deutschland - Ergebnisse einer repräsentativen Studie: Methodik, Alters-, Geschlechts- und Beurteilungseffekte. Zeitschrift für Kinder- und Jugendpsychiatrie, 25, 218 - 233.

Frankel, F., Myatt, R. & Cantwell, D. P. (1995). Training outpatient boys to conform with the social ecology of

popular peers: Effects on parent and teacher ratings.
Journal of Clinical Child Psychology, 24, 300-310.

Inderbitzen, H.M. (1994). Adolescent peer social competence. A critical review of assessment methodologies and instruments. In T.H. Ollendick & R.J. Prinz (Eds.) Advances in clinical child psychology, Vol.16, (pp.227-259). New York: Plenum Press.

Kane, J.S. & Lawler, E.E., III. (1978). Methods of Peer Assessment. Psychological Bulletin, 85, 555-586.

Lefkowitz, M.M. & Tesiny, E.P. (1980). Assessment of Childhood Depression. Journal of Consulting and Clinical Psychology, 48, 43-50.

Lindeman M, Harakka T, Keltikangas-Järvinen L (1997). Age and gender differences in adolescents' reactions to conflict situations: Aggression, prosociality, and withdrawal. Journal of Youth and Adolescence, 26, 339-351.

Lutz, R. (1989). Zur Vorhersagbarkeit von Skalen-Interkorrelationen. Fachbereich Psychologie der Philipps-Universität Marburg (Bericht Nr. 98).

Merrell, K.W. (2001). Assessment of Children's Social Skills: Recent Developments, Best Practices and New Directions. Exceptionality, 9, 3 - 18.

Newcomb, A. F., Bukowski, W. M., & Pattee, L. (1993). Children's peer relations: A meta-analytic review of popular, rejected, neglected, controversial, and average sociometric status. Psychological Bulletin, 113, 99 - 128.

Perry, D. G., Kusel, S. J. & Perry, L. C. (1988). Victims of peer aggression. Developmental Psychology, 24, 807 - 814.

Prinz, R.J., Swan, G., Liebert, D., Weintraub, S. & Neale, J.M. (1978). ASSESS: Adjustment scales for sociometric evaluation of secondary-school students. Journal of Abnormal Child Psychology, 6, 493-501.

Realmuto, G. M., August, G. J., Seiler, J. D., & Pessoa-Brandao, L. (1997). Peer assessment of social reputation in community samples of disruptive and nondisruptive children: Utility of the Revised Class Play method. Journal of Clinical Child Psychology, 26, 67 - 76.

Remschmidt, H. & Walter, R. (1990). Psychische Auffälligkeiten bei Schulkindern. Zeitschrift für Kinder- und Jugendpsychiatrie, 18, 121 - 132.

Reynolds, W.M. (1998). Depression. In T.H. Ollendick (Ed.) Comprehensive Clinical Psychology, Vol.5. Children and adolescents: Clinical formulation and treatment (pp. 418-453). Amsterdam: Elsevier.

Scheithauer, H., Hayer, T. & Petermann, F. (2003). Bulling unter Schülern: Erscheinungsformen, Risikobedingungen und Interventionskonzepte. Göttingen: Hogrefe.

Tesiny, E.P. & Lefkowitz, M.M. (1982). Childhood Depression: A 6-month follow-up study. Journal of Consulting and Clinical Psychology, 50, 778-780.

Autorenhinweis

Patrick Pössel, Universität Tübingen

Ulrike von Delleman, Universität Tübingen

Martin Hautzinger, Universität Tübingen

Danksagung: Bedanken möchten wir uns besonders bei den teilnehmenden Schülern, ihren Lehrern und Eltern, sowie dem Oberschulamt Tübingen, das durch seine Genehmigung die Durchführung der Studie erst ermöglichte.

Patrick Pössel

Eberhard-Karls-Universität

Psychologisches Institut

Abt. Klinische und

Entwicklungspsychologie

72072 Tübingen

e-mail: patrick.poessel@uni-tuebingen.de

Tabelle 1

Rohwerte und korrigierte Trennschärfe der Items, der
Endversion der ASSESS-D in Originalnummerierung (n = 291)

Item	<u>M</u>	<u>SD</u>	<u>rit</u>
01. Welche Schüler verhalten sich unverschämt den Lehren gegenüber?	.11	.19	.85
02. Welche Schüler lügen, um ihre Ziele zu erreichen?	.09	.12	.67
03. Welche Schüler verhalten sich unausstehlich?	.12	.13	.67
04. Welche Schüler beleidigen anderen?	.19	.22	.88
05. Welche Schüler stoßen andere Leute herum?	.10	.17	.86
06. Welche Schüler verhalten sich wirklich wild?	.11	.17	.88
07. Welche Schüler wollen, dass alles auf ihre Art gemacht wird?	.12	.14	.63
08. Welche Schüler streiten sich oft und werden schnell wütend?	.11	.15	.84
09. Welche Schüler nehmen Sachen, die ihnen nicht gehören?	.08	.13	.83
10. Welche Schüler denken, sie könnten jeden zusammenschlagen?	.09	.16	.83
11. Welche Schüler/innen verhalten sich unecht?	.08	.09	-
12. Welche Schüler werden in gewisser Weise ignoriert?	.11	.18	.87

13. Welche Schüler ziehen es vor, allein zu sein?	.05	.11	.84
14. Welche Schüler scheinen nie Spaß zu haben?	.07	.12	.86

Item	<u>M</u>	<u>SD</u>	<u>rit</u>
15. Welche Schüler werden nicht gefragt, ob sie an Gruppenaktivitäten teilnehmen wollen?	.11	.17	.85
16. Welche Schüler sind in ihre eigene Welt zurückgezogen?	.06	.11	.84
17. Welche Schüler sind dem anderen Geschlecht gegenüber schüchtern?	.17	.20	.67
18. Über welche Schüler/innen weiß ich nichts?	.22	.16	.44
19. Welche Schüler/innen sind nervös, wenn sie im Unterricht aufgerufen werden?	.11	.12	-.44
20. Welche Schüler/innen fürchten sich nicht, für sich selbst zu sprechen?	.26	.19	-.44
21. Welche Schüler/innen scheinen Selbstzweifel zu haben?	.08	.11	-
22. Welche Schüler fühlen sich beleidigt?	.09	.10	.73
23. Welche Schüler sind sehr leicht verletzbar?	.09	.12	.73
24. Welche Schüler/innen wirken nervös?	.07	.08	-
25. Welche Schüler können gut in einer Gruppe arbeiten?	.33	.19	.84
26. Welche Schüler machen die meisten Dinge gut?	.27	.21	.63
27. Welche Schüler teilen Dinge mit anderen?	.29	.16	.69
28. Welche Schüler bemerken alles, was um sie herum geschieht?	.15	.11	.53
29. Welche Schüler/innen haben Selbstvertrauen?	.33	.19	-

30. Welche Schüler werden in ihrer künftigen Arbeitsstelle wahrscheinlich gut zurechtkommen?	.33	.20	.78
31. Welche Schüler sind gute Verlierer?	.24	.14	.56

Fortsetzung

Item	<u>M</u>	<u>SD</u>	<u>rit</u>
32. Welche Schüler genießen es, mit anderen Leuten zusammen zu sein?	.47	.20	.63
33. Welche Schüler werden von anderen gemocht?	.45	.23	.71
34. Welche Schüler kontrollieren ihre Launen?	.17	.10	.54
35. Welche Schüler sind schlecht in der Schule?	.18	.22	.68
36. Welche Schüler schaffen es nicht, Aufgaben rechtzeitig abzugeben?	.16	.18	.84
37. Welche Schüler verschwenden Zeit und haben Schwierigkeiten, Dinge zu Ende zu bringen?	.13	.13	.84
38. Welche Schüler verpfuschen gewöhnlich eine Aufgabe?	.12	.13	.82
39. Welchen Schülern ist es egal, wie gut sie etwas können?	.14	.13	.64
40. Welche Schüler geben auf, wenn es schwer wird?	.15	.12	.69
41. Welche Schüler sind unordentlich?	.18	.19	.79

Anmerkungen: M = Mittelwert; SD = Standardabweichung; rit = part-whole korrigierter Trennschärfekoeffizient bezogen auf die jeweilige Skala, der das Item zugeordnet ist; - = Nicht ermittelte Trennschärfe.

Fortsetzung

Tabelle 2

Faktorenladungsmatrize des ASSESS-D in Originalnummerierung
inklusive der Zuordnung der Items zu den Originalskalen

Item	Faktorladungen						h ²
	Origini- nal- skala	Ag- gres- sion	Rück- zug	Emp- find- samkeit	Soziale Kompe- tenz	Schu- lische Probleme	
01	Ag	.77	–	–	–	.37	.78
02	Ag	.52	–	.34	–	.44	.75
03	Ag	.63	–	.31	–	–	.71
04	Ag	.82	–	–	–	.31	.85
05	Ag	.89	–	–	–	–	.88
06	Ag	.89	–	–	–	–	.88
07	Ag	.66	–	.39	–	–	.72
08	Ag	.86	–	–	–	–	.78
09	Ag	.83	–	–	–	–	.78
10	Ag	.85	–	–	–	–	.82
12	Rü	–	.84	–	–	–	.85
13	Rü	–	.91	–	–	–	.83
14	Rü	–	.89	–	–	–	.82
15	Rü	–	.84	–	–	–	.84
16	Rü	–	.90	–	–	–	.83
17	Rü	–	.70	–	–	–	.72
22	An	–	.37	.76	–	–	.73
23	An	–	.42	.71	–	–	.76

Anmerkungen: Ag = Aggression; Rü = Rückzug; An = Angst; So = Soziale Kompetenz; Sc = Schulische Schwierigkeiten. Die Faktorenladungen, nach denen die Items den Faktoren zugeordnet werden, sind fett markiert, - = Ladungen < .30; h^2 = Kommunalität

Fortsetzung

Tabelle 3

Interne Konsistenzen, Interkorrelation der Skalen und
Retest-Reliabilitäten

	Aggres sion	Rück- zug	Empfind- samkeit	Soziale Kompe- tenz	Schulische Probleme
Aggression	.94**				
Rückzug	-.20*	.95**			
Empfind- samkeit	-.12*	.57**	.84**		
Soziale Kompetenz	-.11	-.32**	-.15**	.89**	
Schulische Probleme	.63**	.04	.04	-.32**	.90**
Interne Konsistenz	.95	.91	.84	.89	.91

Anmerkungen: Die Retest-Reliabilitäten sind in der
Hauptdiagonalen dargestellt und fett markiert; * $p < .05$,
** $p < .01$

Tabelle 4

Korrelationsmatrix der ASSESS-D Subskalen mit den Skalen
des YSR als Maßen der Kriteriumsvalidität (n = 291)

	Aggres- sion	Rück- zug	Empfind- samkeit	Soziale Kompetenz	Schulische Probleme
<u>Youth Self-Report</u>					
Externalisierende Störung	.31**	-	-.11	-.04	.30**
Internalisierende Störung	.04	-.02	.11	.07	.01
Andere Probleme	.14*	-.08	.05	-.03	.25**

* $\underline{p} < .05$, ** $\underline{p} < .01$